

Students' Solution of Arrangement Problems and their Connection to Cartesian Product Problems

Erik S. Tillema

Abstract

Two-hour long interviews were conducted with each of 14 sixth grade students, ages 11-12. The purposes of the interviews were to investigate how students solved arrangement problems (APs), and how their solutions of these problems differed from their solution of Cartesian product problems (CPPs). The 14 students represented a balanced mix of students operating with each of three different multiplicative concepts that have been identified in prior research (Hackenberg, 2010). This paper reports on the 11 students who were using the first or second multiplicative concept. Students operating with different multiplicative concepts all experienced similar perturbing elements in their solution of APs relative to their solution of CPPs, but they operated differently to resolve these perturbing elements. These differences are identified and their significance discussed in relation to other research findings on students' combinatorial and multiplicative reasoning.

This is the author's manuscript of the article published in final edited form as:

Tillema, E. S. (2019). Students' solution of arrangement problems and their connection to Cartesian product problems. *Mathematical Thinking and Learning*, 0(0), 1–33. <https://doi.org/10.1080/10986065.2019.1608618>

Introduction

Mathematics educators have provided long-standing recommendations to include enumerative combinatorics in K-12 curricula (e.g., Kapur, 1970). They have made these recommendations because combinatorics problems are closely linked to other important content domains like probability and computer science (Jones, Langrall, & Mooney, 2007; National Council of Teacher of Mathematics [NCTM], 2000; Lockwood, 2013), can foster students' problem solving and posing (English, 1996, 1999; Martin, 2001, Tucker, 2002), and can support students in the development of justification and proof (Maher, Uptegrove & Powell, 2010). These recommendations have impacted K-12 standards (e.g., NCTM, 2000; National Governors Association for Best Practices, Council of Chief State School Officers, 2010), and curricular materials (e.g., Lappan, Fey, Fitzgerald, Friel, & Phillips, 2002; Senk, Viktora, Usiskin, AHbel, Highstone, Witonsky, et al. 1998). Despite the increased presence in curricula and standards, there remains a relatively limited body of research on how students reason about such problems. Moreover, numerous studies have documented that combinatorics problems are challenging for students (e.g., Fischbein & Gazit, 1988; Mulligan & Mitchelmore, 1997), suggesting a need for research that provides models of students' reasoning that can serve as a basis for teachers and researchers to effectively support students in this domain (Lockwood & Swinyard, 2016).

One area of research where models of students' reasoning are particularly important are those that shed light on whether and how students see connections or differences among different kinds of combinatorics problems (Batanero, Navarro-Pelayo, & Godino, 1997; Lockwood, 2011). Author (2014a) provided one example of the way three 8th grade students saw connections and experienced differences between Cartesian product problems (CPPs) like the Outfits Problem and arrangement problems (APs) like the Flag Problem.

The Outfits Problem (CPP): You have four shirts and three pants. An outfit consists of one shirt and one pants. How many possible outfits could you make?

Flag Problem (AP): You have eight colors. How many possible two-striped flags can you make? Assume you count ordered outcomes.

All participants in Author (2014a) were using the most advanced of three multiplicative concepts that have been reported in earlier research (Author, 2009; Hackenberg 2007, 2010; Steffe, 1992, 1994; Ulrich, 2015, 2016), and the findings from this study will be discussed in the Literature Review. The study reported in this paper has a similar focus, students transition from CPPs to APs, but expands the earlier study: the participants in the current study represented a balanced mix of students using all three multiplicative concepts that have been reported in earlier research. The participants of the current study were 14 6th grade students who were asked a sequence of combinatorics problems over the course of two, one-hour interviews, where students started by solving CPPs and then transitioned into solving APs. Given that the earlier study already reported on students using the third, and most advanced, multiplicative concept, this paper focuses on the 11 students who were using the first and second multiplicative concepts. The research questions for this paper are:

- (RQ1) What novel ways of operating were entailed in students' solutions of APs relative to their solutions of CPPs?
- (RQ2) What, if any, qualitative differences were there among students using the same multiplicative concept in their solution of APs?
- (RQ3) What, if any, qualitative differences were there among students using different multiplicative concepts in their solution of APs?

Literature Review

Students' Solutions of CPPs and APs

To date, researchers investigating students' solutions of CPPs have largely done so with elementary grades students (ages 5-11) (e.g., Mulligan & Mitchelmore, 1997; Nunes & Bryant, 1996; Outhred, 1996), while researchers investigating students' solutions of APs have largely done so with middle grades students (ages 12-14) (e.g., Batanero, et al., 1997; Fischbein & Gazit, 1988; see Maher, et al., 2010 for a notable exception). However, one question that neither body of research answers is how students' transition between CPPs and APs. Based on when these problems typically appear in curricula (CPPs in elementary and APs in middle school), studying this transition is important because it marks a transition between combinatorics problems that are appropriate for elementary grades students and those appropriate for middle grades students.

In her investigation of students' solutions of CPPs, English (1991, 1993) identified five strategies that students aged 7-12 used. The strategies ranged from least to most systematic where the least systematic was trial and error, and the most systematic was the odometer method. The odometer method consists of a student holding an item from one set constant (e.g., red shirt), cycling through all of the items of the other set (e.g., blue pants, tan pants, black pants), then changing the item from the first set (e.g., white shirt), cycling through all the items from the second set in the *same order* (e.g., blue pants, tan pants, black pants), and repeating this process until all outcomes are created. Maher and Yankelwitz (2010) have argued that the odometer method is powerful because it can support students to know when, and justify why, they have produced all possible outcomes in their solution of CPPs (see also, Maher & Martino, 1996). For a similar reason, the odometer method is powerful in combinatorics well beyond the solution of CPPs. However, studies with older students indicate that students often do not use the odometer

method in the solution of more advanced combinatorics problems when it would be useful (Eizenberg & Zaslavsky, 2004; Halani, 2013; Lockwood & Gibson, 2016). For this reason, Author (2014a) considered it important to investigate how three 8th grade students (aged 12-13) who used the odometer method to solve CPPs reconstructed this strategy in their solution of APs.

Author (2014a) found that all three students initially solved APs as a sum because they did not count ordered outcomes (e.g., count the red-blue and blue-red flags as two different outcomes). He explained this finding based on three inter-connected issues: (a) the way students used the odometer method in their solution of CPPs; (b) the fact that students' interpreted APs to involve only one set (e.g., the Flag Problem involves 8 colors), while they interpreted CPPs to involve two sets (e.g., the Outfits Problem involves 4 shirts and 3 pants); and (c) the students had not constructed ordered outcomes. The third issue explains why the students in the study did not count ordered outcomes in their solution of APs, but it does not explain how the students' solution of APs as a sum was related to their solution of CPPs.

Author (2014a) explained this relationship by analyzing the students' use of the odometer method in CPPs and accounting for how this method functioned differently in the context of an AP, which the students interpreted as involving only one set rather than two. To solve CPPs using the odometer method, students did not use the first item of one set (e.g., the first shirt) after they had paired it with all items of the other set. Students used this same way of operating to solve APs, like the Flag Problem: They took the first item from the set (e.g., first color) and paired it with all items of the set (e.g., first color, second color, ..., eighth color) to create eight flags. When they got done creating the first eight flags, they did not use the first item again (e.g., first color) because that corresponded to what they had done when they used the odometer method in their solution of CPPs. Doing so left only seven colors from which to make flags, and

so students took the second item from the set (i.e., second color) and paired it with all remaining items of the set (i.e., second color, third color, ... eighth color) to create seven flags. This solution led to producing a sum in the context of solving APs, and the explanation led to an account of how students' solutions of CPPs served as a basis for their solution of APs.

One of the three students in Author's (2014a) study did subsequently solve APs using multiplication. To do so, she first constructed ordered pairs. Once she had constructed ordered pairs she also introduced a novel way of operating. Recall that the students interpreted APs as involving only a single set (e.g., 8 colors) whereas they interpreted CPPs as involving two sets (e.g., 4 shirts and 3 pants). This difference meant that for CPPs the students did not create a second set from the first set before they made outcomes. On the other hand, the students had to figure out how to create a second set in their solution of APs. The student who solved APs using multiplication learned to create a second set from the first where she considered the second set to be *identical to* the first, and established the entire second set *prior* to making any outcomes. This way of operating meant that prior to creating, for example, any flags in the Flag Problem she established that there would be two sets of eight colors—one set of colors for the first stripe and one set of colors for the second stripe. Moreover, she created the second set of colors from the first, and considered the two sets to be identical. The fact that they were identical to each other meant that, for example, if red was the third color in the first set it was also the third color in the second set. This relationship between the two sets meant that she re-constructed the odometer method in a novel way; the order of the items she used in the second position was *the same as, and dependent on*, the order of the items she used in the first position when creating flags. Figure 1 is a list for the Flag Problem that shows this issue: the order of the colors in the second position (highlighted in blue in the first column) is the same as the order of the colors in the first position

(highlighted in red).¹ This feature was neither explicit nor present when she used the odometer method in her solution of CPPs because she considered the order of the elements of one of the sets (e.g., shirts) to be *unrelated to* the order of the elements of the other set (e.g., pants).

G, G	Y, G	R, G	Pnk, G	Bk, G	O, G	P, G	B, G
G, Y	Y, Y	R, Y	Pnk, Y	Bk, Y	O, Y	P, Y	B, Y
G, R	Y, R	R, R	Pnk, R	Bk, R	O, R	P, R	B, R
G, Pnk	Y, Pnk	R, Pnk	Pnk, Pnk	Bk, Pnk	O, Pnk	P, Pnk	B, Pnk
G, Bk	Y, Bk	R, Bk	Pnk, Bk	Bk, Bk	O, Bk	P, Bk	B, Bk
G, O	Y, O	R, O	Pnk, O	Bk, O	O, O	P, O	B, O
G, P	Y, P	R, P	Pnk, P	Bk, P	O, P	P, P	B, P
G, B	Y, B	R, B	Pnk, B	Bk, B	O, B	O, B	B, B

Figure 1. List for the Flag Problem

Author's (2014a) findings were compatible with prior research that had identified middle grades students tended to solve combination and arrangement problems as a sum first (e.g., Fischbein & Gazit, 1988; Piaget & Inhelder, 1975). However, his findings extended these results by providing an explanatory model for: why students would solve APs as a sum based on how they solved CPPs; and what new ways of operating students introduced when they solved APs relative to their solutions of CPPs. Three limitations of Author's (2014a) study were that: (a) only a single student constructed ordered pairs, which did not provide a significant amount of data on the process by which students make this construction; (b) only one of the students considered APs to involve multiplication; and (c) all three students were using the most advanced of three multiplicative concepts, as reported in prior research (Hackenberg, 2007, 2010; Steffe, 1992, 1994), which had ramifications for the way the students solved both CPPs and APs (Author, 2013, 2014). The current study's intent was to address these limitations.

Conceptual Framework

¹ "G" represents the color green; "Y" represents the color yellow; "R" represents the color red; "Pnk" represents the color pink; "Bk" represents the color black; "O" represents the color orange; "P" represents the color purple; "B" represents the color blue.

Product of Measures Problems

Area, volume, and combinatorics problems (including CPPs and APs) are part of a class of problems that Vergnaud (1983) called product of measures problems. One unique characteristic of product of measures problems is the unit that gets counted, for example, an area unit or an outfit, is formed from the multiplicative composition of two more elementary units; to create an area unit a student can multiplicatively compose two length units or to create an outfit a student can multiplicatively compose one shirt and one pants (Simon & Blume, 1994). Behr, et al. (1994) specified how this property meant that the units involved in product of measures problems differ from the units involved in one of Vergnaud's other problem categories, isomorphism of measures problems. The Donut Problem is an isomorphism of measures problem, and can be used to illustrate this difference.

The Donut Problem. You have 3 packages of 4 donuts. How many total donuts are there?

In the Donut Problem students can make a one to many coordination between one package and four donuts to create a package that contains four donuts (Figure 2a). The units (i.e., donuts) inside the containing unit (i.e., the package) are units of one. This unit structure differs from problems like the Outfits Problem in that student may first create the outfits by multiplicatively composing pants and shirts. Doing so means that they might pair the first pants with the first shirt, the first pants with the second shirt, the first pants with the third shirt, and the first pants with the fourth shirt, and then make a one to many coordination to create a unit containing all outfits with the first shirt (Figure 2b). Here the four units inside the containing unit are *not* units of one, but instead are *pairs* (cf. Behr, et al., 1994).

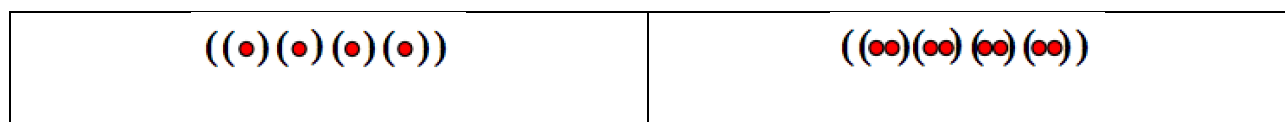


Figure 2a (left) & 2b (right): Difference in unit structures

Author (2016) identified that researchers investigating combinatorial reasoning had not made models of student reasoning that take into account this difference in unit structure. They argued for the need for such models in order to understand how students conceptualized relationships among one and two-dimensional discrete units in the context of solving combinatorics problems (cf. Battista, 2007; Smith & Barrett, 2018). Thus, one of the central goals of Author's (2013, 2014a, in press) prior work has been to define mental operations that students were using that captured whether and how students created a more complex unit structure in their solution of product of measures problems like the Outfits Problem.

Operations, Schemes, Perturbations, and Concepts

Broadly speaking, to characterize students' mathematical reasoning, I use operations, which are mental actions, and schemes, which have three parts, an assimilatory mechanism, an activity, and a result (Piaget, 1970; von Glasersfeld, 1995). The assimilatory mechanism of a scheme involves a student in establishing an interpretation of a problem situation; this interpretation can trigger a particular activity, where the activity of a scheme entails a student in using operations; operations transform a student's initial interpretation of a situation into a result. A student may experience a perturbation when one or more parts of a scheme does not match with his or her expectation. Perturbations can either be consciously conflictive or occur outside of a person's conscious awareness. Perturbations can be occasions for learning because in response a student may change one or more parts of his or her scheme. When the result of a scheme is available to a person prior to operating in a situation, then a person has constructed a concept (von Glasersfeld, 1982). A person's current concepts are what he or she uses to assimilate situations.

Multiplicative Concepts: MC1 and MC2 Students' Solutions of CPPs

Author's (2013, in press) work on students' solutions of CPPs builds on Steffe's (1992, 1994) and Hackenberg's (2007, 2010) work on students' multiplicative reasoning (see also, Ulrich, 2015; 2016). Hackenberg defines three different multiplicative concepts that students use based on the level of embedded units that they can take as a given prior to reasoning about a problem situation. She calls students using the first multiplicative concept, MC1 students; students using the second multiplicative concept, MC2 students; and students using the third multiplicative concept, MC3 students. Steffe (2007) has identified these multiplicative concepts as distinct stages of reasoning that can last for two or more years. This statement, however, does not mean that students do not make modifications in these concepts over this period of time; Steffe and colleagues (1992, 1994; Steffe & Olive, 2010) have provided many examples of the kinds of modifications that students make while using the same multiplicative concept. One modification that researchers using this framework had not addressed was how students might produce a more complex unit structure in their solution of product of measures problems, broadly, and CPPs, specifically. Rather, the multiplicative concepts as outlined by Steffe (1992, 1994) and Hackenberg (2007, 2010) identify a unit structure that is compatible with Vergnaud's (1983) isomorphism of measures problems like the Donut Problem.

To address this issue, Author (2013, in press) identified two operations, pairing and ordering that students used to solve CPPs. A *pairing operation* entails creating a correspondence between two units of one (e.g., a shirt and pants) to create a unit that contains two units, but is counted as a single unit (e.g., an outfit). He termed the units containing two units, *pairs*, in order to differentiate them from units of one. He also defined an *ordering operation* as involving a student in creating an order for the units of a composite unit: for example, a student uses an

ordering operation when she creates a first unit, a second unit, and a third unit in a composite unit of three where evidence for this operation could be assigning a number to each pants as the student solves the Outfits Problem. Author (2013, in press) used pairing and ordering operations along with three operations that Steffe (1992, 1994) previously defined—units coordination, iteration, and disembedding—to characterize students' solutions of CPPs. Here I discuss MC1 and MC2 students' solutions of CPPs, but not MC3 students because MC1 and MC2 students are the focus of data analysis.

Author (in press) found that to successfully solve CPPs using the odometer method, MC1 students assimilate problems like the Outfits Problem using two composite units, for example, three and four. They then order the units of each composite unit, for example, they might order the shirts as, a first shirt, a second shirt, etc., and order the pants as a first pants, a second pants, etc. Once they have ordered the units they can use a pairing operation to create a pair (e.g., the first unit from each composite unit makes an outfit). Typically, MC1 students make each individual pair (e.g., first pants, first shirt; first pants, second shirt, etc.), which means that they iterate their pairing operation, and they often count the pairs they produce by one. These behavioral indicators, however, are not the same across MC1 students especially as they gain experience with this type of problem. The key criteria that is common across MC1 students is to reason about them they have to create them as part of the activity they produce in a situation. One example Author (in press) used to illustrate that MC1 students create pairs in activity was a student who created pairs in the context of making a tree diagram, but then needed to make the pairs again when the interviewer asked her to determine how many outcomes she had represented in her tree diagram; for the student, the pairs were not items she could count outside of the act of making them.

MC1 students can also engage in additional operations when they create pairs; they can use their disembedding operation prior to using their pairing operation. When MC1 students use a disembedding operation, they disembed a unit of one from each composite unit prior to making a pair. The hallmark of a *disembedding operation* is that students establish a pair as *independent from, but related to* the two units of one that comprise it (Figure 3); the importance of a disembedding operation, in this context, is that it allows students to establish a two-dimensional discrete unit, a pair, independently from, but related to, the two one-dimensional discrete units. This issue is shown in Figure 3 by having the pair represented independently from the units of one. When students operate in this way, Author (in press) considers them to have established a multiplicative relationship between a unit of one, a unit of one, and a pair; the reason he considers it *multiplicative* is it is the basis for considering that one times one is equal to one.

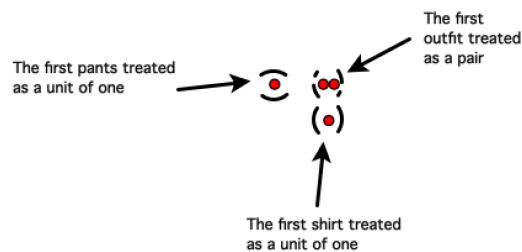


Figure 3. Creating a relationship among a unit of one, a unit of one, and a pair in activity²

Like MC1 students, MC2 students assimilate CPPs using two composite units, for example, three and four, and order the units of each composite unit. They then can disembed one unit from each composite unit and pair them together, creating a multiplicative relationship between a unit of one, a unit of one, and a pair. One difference between MC1 and MC2 students is that MC2 students can *interiorize* this multiplicative relationship. The primary indicator that they have

² In figures like Figure 2, I represent units that a student creates in activity with dashed parenthesis, and units that are interiorized (i.e., not made in activity) with solid parenthesis.

interiorized this relationship is that they can reason with and about pairs even if they have not created them in immediate past experience, which enables them to operate as if one two-dimensional unit is multiplicatively composed of two one-dimensional units.

Some, but not all, MC2 students engage in a units coordination when solving CPPs; Steffe (1992, 1994) has considered a units coordination to involve a student embedding units within units to, for example, create four donuts as one package (recall Figure 2a). In the case of CPPs, this operation involves embedding, for example, four pairs into a containing unit to create four pairs for one pants (recall Figure 2b). Author (in press) illustrated the difference between MC2 students who do and do not make a units coordination through differences in the way they respond to questions like the Extension of the Outfits Problem: Suppose you get one additional shirt (now 5 shirts rather than 4 shirts) and one additional pair of pants (now 4 pants rather than 3 pants). How many *new* outfits could you make? Figure 4 illustrates an array for this problem where the solution to the original problem is represented with red points, and the *new* outfits from the Extension of the Outfits Problem are represented with blue points.

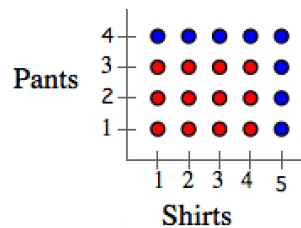


Figure 4. Array for the extension of the Outfits Problem

MC2 students who *do not* engage in a units coordination may predict that: (a) there will be five new outfits, only considering those with the new pants; or (b) there will be four new outfits, only considering those with the new shirt; or (c) there will be a total of nine new outfits, counting the outfit with the new shirt and the new pants twice (the corner point of the array). They can correct these predictions by, for example, filling in the points on their arrays, but the correction is

not straightforward. For example, one student in Author (in press) filled in all new outfits on her array, during which time she filled in the corner point twice, and only realized this issue once the interviewer questioned her. The reason Author (in press) considered these students not to have engaged in a units coordination is they were not yet reasoning with the four pairs (outfits with new shirt) and the five pairs (outfits with new pants) by comparing what was common to each.

MC2 students who *do engage* in a units coordination establish, for example, a multiplicative relationship among a unit of one, a unit of five ordered units, and a unit of five pairs in activity (Figure 5). Establishing this relationship is the basis for seeing a “row” or a “column” of an array multiplicatively where one one-dimensional unit times five one-dimensional units are equal to five two-dimensional units. Creating this relationship enables MC2 students to reason that the outfit with the new shirt and new pants (the corner point) is part of both the outfits with the new shirt and the outfits with the new pants. Therefore, they recognize that it does not need to be counted twice in the absence of working directly with a written list, tree diagram, or array.

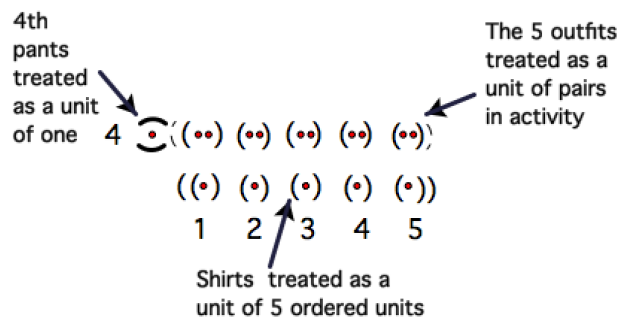


Figure 5. A relationship between a unit of one, a unit of units, and a unit of pairs in activity

The three issues this paper addresses are whether and how MC1 and MC2 students: (a) used the odometer method when solving APs, given that they used this method to solve CPPs; (b) introduced novel ways of operating given that APs can be interpreted as involving only a single set whereas CPPs can be interpreted as involving two; and (c) constructed ordered pairs.

Methods

Interview Study Methodology

A central purpose of semi-structured interviews is to document students' authentic ways of reasoning (Clement, 2000). Like teaching and design experiment methodology (Confrey & LaChance, 2000; Lobato, 2008), semi-structured interviews involve using problems as a context for a researcher to test conjectures about how students might reason. To this end, an interview protocol is developed with conjectures in mind that are framed by current research and a researcher's prior interactions with students. While these conjectures guide the design of the study, a researcher makes modifications to problems as a way to test on-the-spot conjectures that she makes in the context of interacting with students (Steffe & Thompson, 2000). These on-the-spot conjectures allow a researcher to be responsive to unexpected ways of operating that students produce. Unlike teaching and design experiments, it is often not possible to reliably trace out what learning takes place given the small number of interactions in an interview study.

Data Collection and Participants

The school where the study took place was a science magnet school for 6th-12th grade students in a large urban district in the United States. The school population was approximately: 70% African American, 19% Latinx, 7% White, and 4% Multiracial; 65% girls and 35% boys; with 77% of students receiving free or reduced lunch. The district had a school choice model, which meant that to attend the school students from the entire district applied and were selected. To maintain enrollment at the school, students needed to maintain good academic standing and attendance. All 6th grade students in the school enrolled in a course that was designed to prepare them to be ready to take algebra in the 8th grade. Approximately 10% of the student population enrolled in a college level calculus course by the time they were in 12th grade.

The research team consisted of a mathematics education researcher, a graduate student, and two undergraduate researchers. For the first four months of the school year, members of the research team worked one day a week in a 6th grade classroom to familiarize themselves with the students and teacher. This included working individually and in small groups with students, co-teaching or teaching the class, and co-planning with the teacher. The classroom teacher was interested in supporting student thinking but was inexperienced at making this central to her instruction. Thus much of her instruction focused on mathematical procedures, but she was open to suggestions about how to support student thinking in her instruction. Prior to the interviews, the students had no experience with combinatorics problems during their 6th grade. One marker of this lack of familiarity was that 10 of the 11 MC1 and MC2 students initially gave an incorrect answer to the first CPP they solved.

After the first four months of the school year, 16 students volunteered to participate in a selection interview. The selection interview consisted of an un-recorded hour-long interview involving non-combinatorial problems with the goal of identifying students' multiplicative concepts (Appendix A). From the 16 students 14 were asked to participate in the study: three were MC1 (21.4% of the sample), eight were MC2 (57.2% of the sample), and three were MC3 (21.4 % of the sample, not reported on in this paper). Of the 16, two students were not asked to participate because they were also MC2 students; given that the goal was to have a balanced mix of students the research team decided not to include any more MC2 students. Steffe (2007) has estimated that in the general 6th grade population 30% of students are MC1, 30% are MC2, and 40% are MC3. Therefore, the percentage of students' in the study under represented MC1 and MC3 students, and over represented MC2 students. However, the students who participated in the study were representative of the racial and gender demographics of the school.

In the selection interview, the standard method of making a determination of multiplicative concept is to use problems that can involve embedded units (e.g., a package contains 8 candies; a box contains 6 packages; a crate contains 4 boxes; a truck contains 5 crates) (Hackenberg, 2007). Within a given problem a student is asked about relationships they see among units (e.g., How many candies are in a truck?) that are developed to be responsive to the student and test conjectures about whether and how students are embedding units within each other. A comparison of responses across all problems is used to infer multiplicative concept where during an interview the interviewer uses multiple cycles of problem posing in order to test conjectures about students' multiplicative concept.

After the selection interview, the research team conducted two video-recorded, hour-long interviews³ involving combinatorics problems (Appendix B). The first interview was structured to build towards the solution of APs, where students considered ordered outcomes. To accomplish this goal, students were presented first with CPPs whose statement refers to two sets (e.g., three pants and four shirts) where the qualitative referent for the sets was different (e.g., shirts and pants). Then they solved CPPs whose statement referred to two sets where the qualitative referent for the sets was less clear (e.g., You have eight spade cards. I have eight heart cards. How many two-card hands can we make that have one spade and one heart?). The point of presenting problems like the Card Problem was that both people had cards so students had to explicitly introduce a distinction between the suit of the cards. In doing so they had the opportunity to consider how the outcome 3 of hearts-8 of spades was different from the outcome 8 of hearts-3 of spades. The students were then asked an AP involving a small number of outcomes (flipping one coin twice) where they had the opportunity to create most or all possible

³ Given school schedules, five students participated in three forty minute interviews instead of two hour long interviews.

outcomes from their activity of flipping a coin. This provided the opportunity to experientially differentiate between the T-H outcome and the H-T outcome. Then they were asked The Flag Problem, which refers to only one set. This problem sequence was intended to support students to consider ordered outcomes in their solution of APs. All students solved at least the Coin and Flag Problem, and students who moved quickly solved more APs.

The research team anticipated that a likely progression for representations of the problems would be from a picture, to a list, to a tree diagram, and finally to an array, but they also wanted to provide space for students to make sensible decisions about what representation they used. So, for example, the interviewer would ask a student who solved the first problem using a picture questions to support the student to create a list starting with relatively open questions (e.g., I: Can you think of another way to represent the problem?) and moving to more guided questions (e.g., I: What outfit does this picture represent? S: Red shirt and blue pants. I: Can you represent that outfit using letters? S: [writes “RB” or “Red-Blue”] I: Can you represent all of the outfits that way?). Throughout working with students on making representations, the interviewer was careful to observe how students produced representations, to ask probing questions to determine student interpretations of representations, and to have students use multiple representations.

Method of Data Analysis

Data analysis occurred in two phases—ongoing and retrospective analysis (Steffe & Thompson, 2000). On-going analysis included regular meetings of the research team to de-brief and discuss each interview. These meetings included reviewing notes taken during the interviews and watching video of interviews to examine key moments in them. The primary purpose of this phase of analysis was to identify current conjectures about students' reasoning and possible ways to test these conjectures in the context of the interactions. Retrospective analysis occurred after

the interviews were done. The mathematics education researcher transcribed the interview data, wrote low-inference data summaries, and memos that contained conjectures about the data (Corbin & Strauss, 2008; Saldaña, 2013). The research team then met to discuss the transcripts, data summaries, and memos to triangulate interpretations (Mathison, 1988). This process was iterative in that the mathematics education researcher continued cycling through the data, and discussing with the research team until a stable account of each student's reasoning was outlined. To the extent possible (given the small number of interactions) the goal was to make a second-order model of students' reasoning—a researcher's account of the ways and means of operating of a student and any changes to that reasoning (Steffe, von Glasersfeld, Richards & Cobb, 1983).

Data Analysis

Prior to presenting any of the students an AP, all 11 had successfully solved at least two CPPs using the odometer method, and had used each of a list, a tree diagram, and an array at least once. The data analysis of students' solutions of APs focuses on the Flag Problem for two reasons: First, all students worked on this problem; second, it was the first AP students solved where they could not easily produce most or all of the outcomes simply through an experiment (i.e., flipping a coin). The Flag Problem was presented orally to students. The students were given eight colored pencils (red, orange, yellow, green, blue, purple, pink, and black) and a handout that had the outline of nine two striped flags so that they could color in some example flags. While they colored in example flags, they were asked questions like, "Could you have a flag that has red in the top stripe and red in the bottom stripe?" and "What is the difference between the red-yellow flag you made and the yellow-red flag you made?" Once students had colored in example flags on their handout, considering what they might count, they were asked to predict how many possible flags they could make. None of the 11 students made an accurate

prediction. The initial perturbation they experienced was how to operate on the eight colors in order to make flags, which was evident in student comments like, “I split the eight colors into two groups (of four),” where the student predicted there would be 16 flags because four colors paired with four colors would produce 16 flags. For this reason, all students were asked to create a list, a tree diagram, and/or an array. As they worked on the problem the interviewer prompts from the work on their handout along with the structure of the prior problems supported all students to count, for example, the blue-green and green-blue flags as different in their solutions. The data analysis will examine three issues related to students' solutions of APs: a. whether and how students reconstructed the odometer method; b. how students operated differently because they interpreted the problem as involving one set, rather than two; and c. criteria necessary to infer students' construction of ordered pairs and information about how they used ordered pairs to structure the outcomes.

MC1 Students' Solutions of APs

Table 1 provides a summary for the three MC1 students on each issue. To illustrate each, I use data from Carlos's solution of the Flag Problem. As I discuss Carlos's solution, I provide comparison and contrast of his solution to other MC1 students.

Table 1. Summary of MC1 students' solutions of APs

	Odometer Method	Operating with a Single Set	Ordered Pairs
Carlos			Constructed ordered pairs; found ordered pairs in pairs
Darryl	Did not reconstruct odometer method	Separated single composite unit into two parts	Constructed ordered pairs; found ordered pairs in pairs
Alana			Did not construct ordered pairs

Data for issue 1: The odometer method. To start the Flag Problem, Carlos colored in six flags on his handout (blue-green; red-purple; yellow-black; orange-pink; purple-blue; and green-blue), and created a list for these flags (Figure 6).⁴ The interviewer asked if that counted *all*

A handwritten list of six flags: BGr, RR, YE, OPnk, PBl, GB.

Figure 6. Carlos lists the six flags he has colored in on his handout

possible flags, and Carlos began to list flags that had blue in the first position, excluding the blue-blue flag (Figure 7). As he created Figure 7, he visually scanned the colored pencils to

A handwritten list of flags with blue in the first position: BGr, Br, BB, BY, BRR, BPR, BIO.

Figure 7. Carlos lists the flags with blue in the first position

determine that he had used all of the colors except blue in the second position. He then began the second column of his list, which represented all flags that had red in the first position. To create the second column of his list, he used the order of the colors in the first column to help him keep track of which colors he should pair with the color red (Figure 8 shown in red).

He encountered two issues in using the order of the colors from the first column to establish an order for the colors in the second column. The first was that the color in the second position of the first flag in each column was not the same (Figure 8 shown in blue), which meant that he

⁴ He used the letters “Bl” for blue, “B” for black, “P” for purple, “Pnk” for pink, “G” or “Gr” for green, “R” or “r” for red, “Y” for yellow, and “O” for orange.

listed the red purple flag twice in the second column (“RP”) and did not list the red-green flag. The second issue was that the color blue did not appear as a color in the second position in the first column of his list because he had not included the blue-blue flag. He did, however, remember to include the red-blue flag in the second column—he recorded it as the second flag.

BlGr	RP	YB	OPnk	PBl	GBl
Blr	RBl				
BlB	RB				
BlY	RY				
BlPnk	RPnk				
BlP	RP				
BlO	RO				

Figure 8. Carlos writes the second column of his list

To create the third column of his list, he again relied on the prior column to guide the order of the colors for the second position. As he did, he independently noticed both issues and found a way to address them: He skipped the color black when he saw it in the second column of his list since he had already recorded the yellow-black flag as the first flag in the third column (Figure 9, shown in red); he decided to include the yellow-yellow flag because then he did not have to skip the color yellow when he saw it in the second column (Figure 9, shown in blue); and he

BlGr	RP	YB	OPnk	PBl	GBl
Blr	RBl	YBl			
BlB	RB	YY			
BlY	RY	YPnk			
BlPnk	RPnk	YGr			
BlP	RGr	YO			
BlO	RO	Yr			
BlBl	RR	YP			

Figure 9. Carlos finishes the third column of his list

included the yellow-purple flag, recorded as “YP”, as the last flag in the third column because the color purple was in the first flag of the second column and so he knew he had not used purple yet in the third column (Figure 9, shown in purple). He continued to use this pattern over the

course of creating the rest of the columns of his list. Before finishing the third column, he also adjusted the first column to include the blue-blue flag, adjusted the second column to include the red-red flag (Figure 9, shown in orange), and replaced the red purple flag that he had written twice in the second column with the red green flag (Figure 9, shown in green).

Analysis of Issue 1: The odometer method. Carlos's actions indicate that initially he intended to pair, for example blue, with all colors except blue, and subsequently that he did include flags like the blue-blue flag, which he did largely to ensure that he could use his prior column (e.g., second column) of notation to keep track of which colors he had used in the second position of his current column (e.g., third column). He did not, however, follow a fully consistent order for the colors in the second position from column to column. Ultimately, he produced a consistent *pattern* to ensure that he used all colors in the second position in each column, but not the same *order* for the colors, indicating that he did not re-construct the odometer method in the context of solving an AP. Other MC1 students were similar in that they used a consistent pattern for the colors in the second position, but did not use exactly the same order.

Data for Issue 2: Operating with a single set. After Carlos finished his third column of notation, the interviewer asked him how many flags he would list in the fourth column prior to his actually listing them. Carlos responded by saying, "I think the same thing as these," pointing to the first three columns. The interviewer asked him how many flags were in each of the first three columns. Recall that at this point there were eight flags in each column. He indicated he was uncertain and counted the flags in the first column by one, skipping the first flag in the column, and said there were "seven" flags. He then counted the flags in the third column, again skipping the first flag, and said "yeah, seven." The interviewer asked him to recount the flags. This time he counted the first flag in the third column, and said, "Oh, there is eight." He

continued his list, and on several occasions he counted the number of flags in a column, getting eight each time. When he finished his list, the interviewer asked him the following.

Data Excerpt 1: Carlos considers how many flags he could make with the color blue⁵

I: With the blue one, how do you know how many flags you can make with the blue?

C: You just do all the colors that you could do. Like blue and yellow, blue and black, blue and red, blue and purple. You are going to use all the colors.

I: All the colors. Is there a way you could figure out how many colors there are to use with blue? In other words, how many you are going to make without writing them all out?

C: You could count up the colors there are and so there is seven colors. [He does not count the colored pencils].

I [Surprised because this response does not match with the list he has created]: There are seven colors you could use with blue? [Carlos nods his head “yes.” The interviewer continues to question him, and, after several rounds of questions, Carlos does explain why there could be eight flags with the color blue.]

Analysis of Issue 2: Operating with a single set. Carlos's response that there were “seven colors” that he could use with the color blue was surprising. One interpretation of his response is that he simply thought there were a total of seven colored pencils, including the color blue, rather than eight. I don't make this interpretation given features of his solution, together with the way that other MC1 students responded to solving the Flag Problem. For example, Darryl, another MC1 student, responded to a similar question by saying, “...so if you have orange on top and you have eight colors you have to subtract it from one is seven.” These types of responses were

⁵ In the Data Excerpts, I is for the interview-researcher, W is for a witness-researcher, and all other letters stand for the student. Comments enclosed in brackets [] describe students' nonverbal action or interaction from the interviewer's perspective. Ellipses (...) indicate a sentence or idea that seems to trail off; four periods (....) denote omitted dialogue or interaction; and parenthesis () are the researcher's interpretations of a word meaning.

typical for MC1 students and explain why they did not, at least initially, count flags like the blue-blue flag; they thought about removing one color from the total number of colors to put in the first position, and then had seven other colors remaining. Carlos's actions and response illuminate that this issue was not a superficial one because even after he included flags like the blue-blue flag, he still seemed to interpret the situation as involving one color, the color blue, and seven colors that could go with the color blue. He did eventually identify that eight colors could go with the color blue, but he did so in response to multiple questions from the interviewer.

My explanation for this way of operating is that MC1 students assimilated APs using a single composite unit (Figure 10a). Carlos, then had to figure out how to operate on the single composite unit in order to create pairs; he operated on it by separating it into two parts, one unit (the color blue) and seven units (the seven remaining colors) (Figure 10b). I make the inference



Figure 10a (left), 10b (right). Carlos assimilates with one composite unit and separates it that he separated the composite unit into two parts rather than *disembedded* a unit of one from the composite unit. The key difference in these two interpretations is that if he had used a disembedding operation he would have established the unit of one as *independent from, but related to* the composite unit of eight; it would have been one of eight colors that could be used to create flags even if ultimately he decided not to use the color blue to create a blue-blue flag. The language that MC1 students used (e.g., Darryl's language "you have to subtract" the one color) was much more consistent with separating the original colors into one color for the first position and seven colors for the second position. This interpretation means there is a need to

explain Carlos's inclusion of flags like the blue-blue flag; I interpret his inclusion of this kind of flag as an outgrowth of his using the prior column in his list to keep track of what colors he had already used in the current column. However, I don't infer that including this kind of flag indicated that he changed how he was operating, an inference that is supported by his statement after he finished his list that seven colors could go with the color blue.

Once Carlos separated the composite unit I infer he used his pairing operation. As he made pairs, a key question was whether he retained a composite unit of *seven* as part of the situation. Here recall that the interpretation is that he was not working with one color and eight colors, but rather one color and seven colors, even though he did list eight flags in each column. I infer that he could *recall* that there were seven colors for the second position after he made flags in a single column, but as he was making the flags the seven colors were "used up" or "exhausted." I base this inference on the fact that when he was asked how many flags were in each of the first three columns he expressed uncertainty, and then counted the number of flags in the first and third column. Had he retained the composite unit of seven as he was making the pairs I infer that he would have likely just stated that there would be seven flags in each column prior to counting. Figure 11 shows this issue through the removal of the first unit, the color green, from the composite unit once the first pair, the blue-green flag, is created. From the perspective of operations, Figure 11 is intended to show that he *did not* use a disembedding operation as he was creating pairs. Using a disembedding operation would mean that he produced the pairs as independent from, but related to, the two units of one that comprised it.

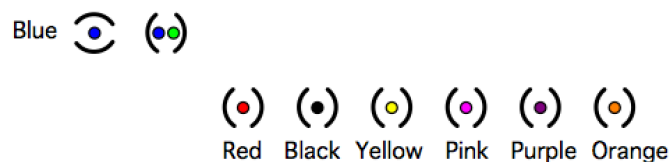


Figure 11. Carlos creates pairs

It is interesting that when he did count the total number of flags in the first and third columns he initially counted only seven flags; he counted all eight only after interviewer prompting. Doing so provides supporting evidence for my interpretation that outside the immediate experience of making flags he could *recall* that he separated the composite unit of eight into one and seven. It supports this interpretation because I infer he could recall separating or actually did separate the composite unit again prior to counting, which provided him with some expectation that he would create seven flags in the first and third column prior to his counting. This expectation led to his counting only seven flags in each column. However, I do not interpret his counting seven flags to mean that he maintained the seven colors while he was actually making flags because he expressed uncertainty about the number of flags in each column.

This interpretation means I considered Carlos to be creating pairs in activity—where the pairs he produced constituted his “mathematical world,” *but not* to be producing a multiplicative relationship among a unit of one, a unit of one, and a pair where he established the pair as independent from, but related to the units of one that comprised them. Carlos’s way of operating was representative of the other MC1 students. However, it was the most complicated response among them because of his inclusion of flags like the blue-blue flag. The other MC1 students simply did not consider these as possible flags to be counted, and in Darryl’s case stated that they were “removing” one color from the eight colors to establish one color and seven colors.

Data for issue 3: Construction and use of ordered pairs. Carlos’s solution also highlights the criteria I used for identifying students’ initial construction of ordered pairs.

Data Excerpt 2: Carlos establishes ordered pairs in the Flag Problem⁶

⁶ Portions of this data excerpt occurred prior to the data that was shared in analyzing issue 1 and issue 2.

I [Carlos has colored in five flags on his handout—blue-green; red-purple; yellow-black; orange-pink; purple-black. The fifth flag has purple in the bottom stripe and the first flag has purple in the top stripe]: So here ah purple is in the top stripe and here purple is in the bottom stripe. Does that make any difference?

C: No. [Carlos pauses for three seconds.] It would make a difference if I would have green and blue. [Carlos points to the blue-green flag he has colored on the handout].

I: Oh, would you, could you actually make that flag?

C: Yeah [colors in a green-blue flag on the handout.]

I: Okay so if we were going to count all of the different flags that you could make if you just kept on doing all the flags. Do you want to count these two as different or the same [points to the blue-green and green-blue flag]?

C: Well they would be different because in one the green stripe is here at the bottom instead of the top.

....

I [Carlos has completed his list]: How can you tell the difference between these when you wrote them down [points to “BIG”, and “GBI” in his list]?

C: Because blue comes first in this one and blue comes last in this one.

....

I: So if you have the blue, well first of all, remind me again which is the top stripe and which is the bottom stripe. Would you label that in your picture (meaning in his list)?

C: Yeah I am using like this example. [He writes “BIG”, and then writes top under “BI” and bottom under “Gr”]. Like the second one would be the bottom.

Analysis of issue 3: Construction and use of ordered pairs. Carlos introduced the difference between a green-blue and a blue-green flag in response to the interviewer's question about whether it mattered if the color purple was located in the top or bottom stripe, and decided that he would count these two flags as different because the "green stripe is here at the bottom instead of the top." Doing so was necessary, but not sufficient, evidence to conclude that he had established ordered pairs; it was necessary because it meant he established that the two flags could be counted as different, but it was not sufficient because, with only this evidence, the difference in the two flags could have simply been based on the fact that they looked different visually. The second criteria that I considered essential to infer the construction of ordered pairs was Carlos differentiated between the blue-green and green-blue pair based on the position of each color in the pair, saying, "because blue comes first in this one and blue comes last in this one," referring to his written notation for the two pairs. The third criteria I considered essential to infer the construction of ordered pairs was that he explicitly connected each position in the pair to one of the stripes in the flag—"Like the second one (the color in the second position) would be the bottom (stripe)." I considered these criteria essential to impute the construction of ordered pairs to students because it indicated that they considered the flags to be different; they related this difference to the position in which a color appeared in the pair; and they related the position of the color in the pair back to the original difference that they established (first position means top stripe and second position means bottom stripe).

Of the three MC1 students, two provided evidence of all criteria essential to the construction of ordered pairs. The third MC1 student considered flags like the blue-green and green-blue to be different from one another based on the colors being in the top or bottom stripe (first criteria), but did not provide evidence of establishing the second or third criteria.

MC2 Students' Solutions of APs

Table 2 provides a brief summary for the eight MC2 students on each of the three issues that I analyze. To illustrate each issue, I use data from three students' solutions of the Flag Problem. As I discuss each issue, I provide comparison and contrast for all MC2 students on that issue.

Table 2. Summary of MC2 students' solutions of APs

	Odometer Method	Operating with a Single Set	Ordered Pairs
Shonda	Reconstructed the odometer method.	Disembedded a unit of one from the composite unit, disembedded pairs, evidence of creating a multiplicative relationship among a unit of one, a unit of eight units, and a unit of eight pairs.	Constructed ordered pairs; used ordered pairs to find a row and column in list/array.
Terionna			
Tiana			
Jose			
Deondria		Disembedded a unit of one from the composite unit, disembedded pairs, evidence of creating a multiplicative relationship among a unit of one, a unit of one, and a pair.	Constructed ordered pairs; found ordered pairs in pairs.
Jada			
Bernice	Did not reconstruct the odometer method.		Did not construct ordered pairs.
Kai	N/A did not solve an AP successfully	N/A, did not solve an AP successfully.	N/A, did not solve an AP successfully.

Data for issue 1: The odometer method. Jada began the Flag Problem by coloring in all nine flags on her handout. The interviewer asked if those were all of the possible flags to which she responded, “no,” and so the interviewer asked her to make a list, a tree diagram, or an array. Figure 12a shows the beginning of her list. After she completed the second column, the interviewer asked her, “Did you use the same order here [pointing to the colors in the second position of the first column] and here [pointing to the colors in the second position of the second column]?” Jada said “yes,” stating that the order of the colors in the second position was “green, yellow, red, pink, black, orange, purple, and blue.” She then completed her list (Figure 12b).

B-G	G-G	B-G	G-G	Pink-G	Y-G	R-G	O-G	Blue-G	Purple-G
Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R	R	R	R	R	R	R	R	R	R
P	P	P	P	P	P	P	P	P	P
B	B	B	B	B	B	B	B	B	B
O	O	O	O	O	O	O	O	O	O
P	P	P	P	P	P	P	P	P	P
B	Blue	B	B	B	B	B	B	B	B

Figure 12a (left) & 12b (right). Jada's list

Analysis of issue 1: Odometer method. Jada's list and statement indicates that she used the same order for the colors in the second position from column to column (see Figure 12b in blue). She, however, did not use the same order for the colors in the first position as she had for the colors in the second position. That is, the order of the colors going across (Figure 12b in red) was different from the order of the colors going down (Figure 12b in blue). This way of operating indicated that Jada treated the order of the colors in the first and second position as *independent* of one another. Six of the eight MC2 students operated similarly to Jada when they produced lists or tree diagrams in their solution of APs; they used a consistent order for the colors in the second position from column to column, but this order was not the same as the order for the colors they used in the first position. One of the other two MC2 students was like the MC1 students in that she did not use the same order for the colors in the second position from column to column, and the other MC2 student did not successfully solve an AP.

Data for issue 2: Operating with a single set. Shonda's solution of the Flag Problem illustrates well how MC2 students operated when they had only a single set out of which to make outcomes. Her solution is also a good contrast to the MC1 students because she initially seemed to operate similarly to the MC1 students, but then changed how she was operating. To solve the Flag Problem, Shonda simultaneously made a tree diagram and a list.

Data Excerpt 3: Shonda's list and tree diagram for the Flag Problem

S [After coloring in six flags on her handout, the interviewer asked Shonda to predict the total number of flags. She said she was uncertain, and so the interviewer asked her what she would do to figure out the total number of possible flags]: Like blue, blue could be matched up with [counts the first four colored pencils on the table quietly] five, six, seven different colors. And that'd be orange, pink, purple, green, yellow, black, red. [Shonda makes a tree diagram first (Figure 13a). To make her tree diagram she uses the “B” in the center to mean blue, and the “B” on the leaf to mean black. She then uses Figure 13a to make Figure 13b, eliminating the “BB” flag when she makes Figure 13b presumably because she thinks it is the blue-blue flag.]

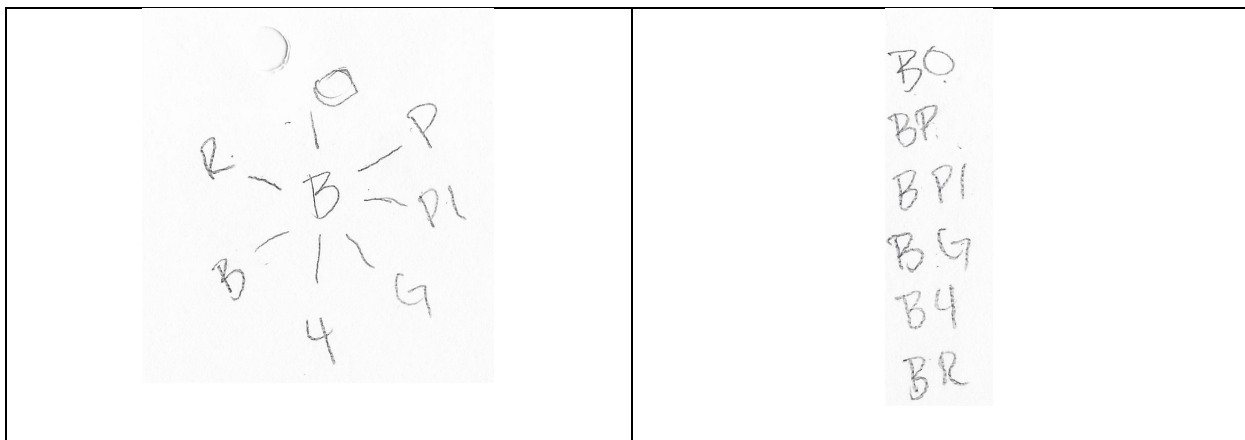


Figure 13a (left) & Figure 13b (right). The beginning of Shonda's tree diagram and list.

....

S [sees the “B-B” flag in her tree diagram]: Wait, but then there could be a blue-blue (flag).

I: A blue-blue (flag)? How many did you make in this one [points to her tree diagram (Figure 13a)]? Is this a blue-blue (flag) [points to Figure 13a where the B-B combination is recorded]?

S: Is it? Wait, that is black. Okay. My bad. So that is black [writes “k” after the B in Figure 13a].

She then adds the blue-blue flag to her tree diagram (Figure 14a) and the blue-blue and blue-black flag to her list (Figure 14b).]

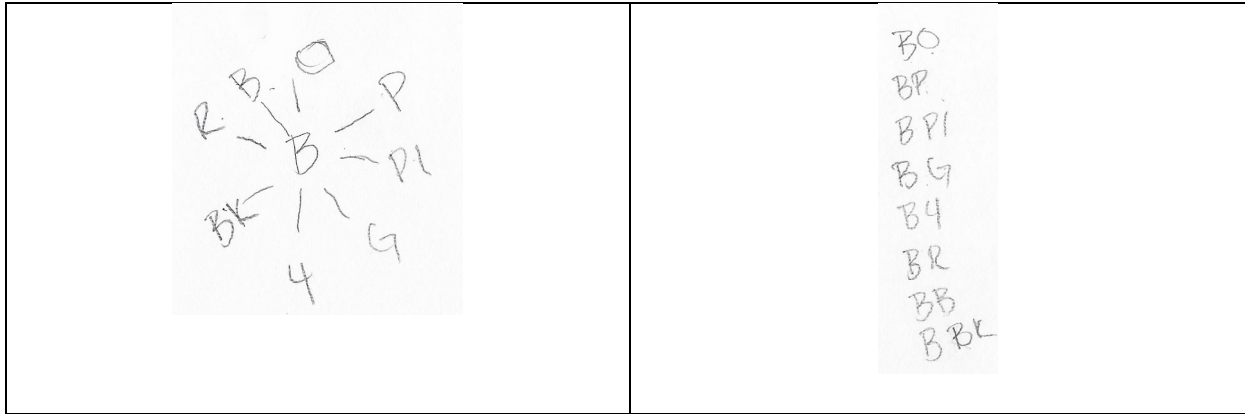


Figure 14a (left) & Figure 14b (right): Shonda adjusts her tree diagram and list

....

I: And how many did you make [points at Figure 14a] on the tree?

S [counts the number of flags in Figure 14a]: Eight, wait...[She appears surprised. She re-counts the number of flags in Figure 14a. She says confidently,] Okay so it is eight combinations you could make. [Prior to continuing her list and tree diagram, she adds a “B” to the upper left corner of her list to show that this column is all the flags made with the color blue in the first position (Figure 15, underlined). She continues her tree diagram and list. Each time she starts a new column in her list she first writes a letter like the underlined “B” in Figure 15 before writing the flags she could make with that color.]

Figure 15 shows the same list as Figure 14b, but with the first column header 'B' underlined. The list is: B BO, BP, BPI, BG, B4, BR, BB, BBL.

Figure 15. Shonda adjusts her list to show that this column uses the color blue

....

I [Shonda has completed her tree diagram and list except for the flags that have black in the first position. She has not counted the number of flags she could make with a single color since she established that she could make eight flags with the color blue. The interviewer starts to ask her how many flags she thinks she will make with the color black]: Do you have a prediction, how many you think...

S [cutting the interviewer off]: Eight. [Both she and the interviewer laugh].

I: You don't want to make it?

S: Well I can. [Shonda indicates she considers listing the final flags unnecessary.]

Analysis of issue 2: Operating with a single set. To start her solution of the Flag Problem, I infer Shonda operated similarly to MC1 students. She referenced that there would be seven colors that could go with the color blue (Lines 671-672), indicating that she assimilated the situation using a single composite unit of eight and separated the composite unit into one and seven. However, there was evidence that she changed how she operated as she made her tree diagram and list. One piece of evidence was her inclusion of the blue-blue flag, which was, in part, a happy accident in that she momentarily interpreted the "B-B" flag in her tree diagram as the blue-blue flag (Line 679), even though she seemed to intend it to be the blue-black flag (Lines 671-673, Lines 680-682). Her inclusion of the blue-blue flag differed from Carlos; Carlos had primarily included flags like the blue-blue flag in order to ensure that he could use the previous column of his list to help him keep track of what flags to make in the current column. Shonda, however, seemed to include it because she considered it to be no different than any of the other flags that had blue in the first position, stating, "wait, but then there could be a blue-blue (flag)." She also differed from Carlos in that it seemed to make sense to her that if she included blue as one of the colors for the second position she would get eight flags when blue

was in the first position, stating, “Okay so it is eight combinations you could make.” Some MC2 students, for example, Jose expressed this relationship verbally as, “So (it’d be) eight (flags for one color) because one times eight equals eight,” a multiplicative statement that referred to one color, eight colors, and eight flags.

These differences led me to a different interpretation of the operations Shonda, and the other MC2 students, produced to solve APs. I infer that Shonda assimilated the situation with a single composite unit of eight (Figure 16a), and disembedded a unit of one from the composite unit (Figure 16b). Using a disembedding operation established a unit of one, independent from, but related to the composite unit of eight. This interpretation is consistent with her inclusion of the underlined letter “B” in Figure 15; she explicitly recorded the one color that she was using to make all of the flags in that column, independently of the flags she listed. I infer she, then, used

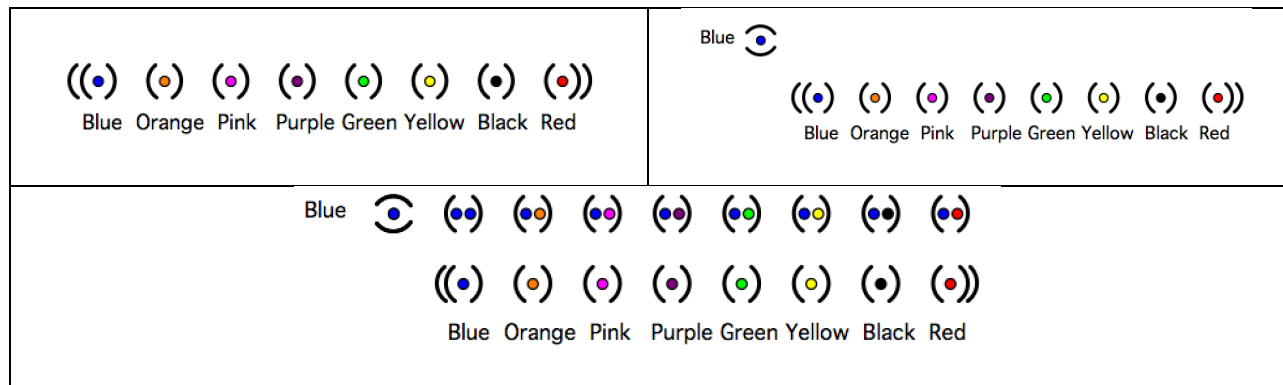


Figure 16a (top left), 16b (top right), & 16c (bottom). Shonda's activity

her disembedding operation again in concert with her pairing operation to establish the pairs independently from, but related to the composite unit of eight (Figure 16c). One piece of evidence that she established the pairs independently from, but related to, the one color and eight colors was her statement that she would make eight flags with the color black prior to actually listing them. This statement indicated that she had adjusted her expectation of how many flags she could make with each individual color, even though she had not counted the number of flags

she could make with any of the colors except blue. Another piece of evidence were statements like Jose's that "one times eight was equal to eight," indicating that the one color, the eight colors, and the eight flags were all independent of each other.

Shonda's way of operating was representative of seven of the eight MC2 students; the eighth student did not correctly solve the problem. The seven students who did operate in this way provided evidence that not all of them established the same multiplicative relationships in these situations. This difference is illustrated through differences in how they constructed and used ordered pairs to structure the outcomes, the final issue that I analyze.

Data for issue 3: Construction and use of ordered pairs. Recall the criteria I considered essential to infer the construction of ordered pairs: (a) a student considered flags like the blue-green and green-blue flag to be different; (b) related this difference to the position in which a color appeared in a pair (e.g., green is in the first position); and (c) related the position back to the original difference that they established (e.g., first position means top stripe). I could attribute these three criteria to six of the eight MC2 students; the other two established the first criteria, but did not establish the second two criteria. For the MC2 students who constructed ordered pairs, I was interested in determining whether they could also see that all ordered pairs with a particular color (e.g., black) made a row and column in their lists or arrays (Figure 17).⁷ The

G,G	Y,G	R,G	Pnk,G	Bk,G	O,G	P,G	B,G
G,Y		...		Bk,Y		...	
G,R				Bk,R			
G,Pnk		...		Bk,Pnk		...	
G,Bk	Y,Bk	R,Bk	Pnk,Bk	Bk,Bk	O,Bk	P,Bk	B,Bk
G,O		...		Bk,O		...	
G,P				Bk,P			
G,B		...		Bk,B		...	

Figure 17. A list showing all flags with black in either the first or second position or both

⁷ This problem has similarities to the Extension of the Outfits Problem outlined in the Theoretical Framework. However, this problem is different in two ways: a. no new colors are added; and b. a single color, rather than one shirt and one pants, creates the row and column structure. Recall from the Theoretical Framework that some MC2 students could solve the Extension of the Outfits Problem and others could not.

following data excerpt with Terionna illustrates this issue.

Data Excerpt 4: Terionna uses ordered pairs to structure the outcomes⁸

I [Terionna has represented the flags using a list, tree diagram, and array. The interviewer covers all representations she has made for the problem]: Can I ask you a challenge question? [Terionna nods her head “yes”]. Do you know how many total flags you could make that would have black in them in some way? Either in the top stripe or the bottom or in both?

T: Maybe eight. Eight different kinds of combinations. [The interviewer asks Terionna to verbally list the eight flags. She lists them starting with the red-black flag and ending with the black-black flag.]

....

T: Wait, you [“you” refers to herself not to the interviewer] are just making the black on the bottom, and then the other colors on top.

I: Yeah that is right. So how many total would have black in either stripe?

T: Sixteen. Yeah, sixteen.

I: Okay so go ahead and tell me why it is sixteen. That is good.

T: Because, well, you said how many could have either black on top or black on bottom.

There are eight and you could put it on, put the eight different colors on top of the black, and then you could put it on the bottom, if you switched it around, and if you put it altogether, it would be sixteen different groups of the black first and color, and then color and then black second. [The interviewer asks her to verbally list the sixteen flags.

She lists the first eight flags, saying the black-black flag as the eighth flag. She then

⁸ We note that Terionna uses an array in her solution to the problem. Students' transition to arrays entailed some changes in their reasoning. We do not analyze those changes here because the reasoning that Terionna produced largely occurred prior to her consideration of the implications it had for her array representation.

pauses for ten seconds, and then says]: So it wouldn't be [Terionna means it would not be sixteen]. So black and black wouldn't be one group again plus another group because you can't switch black and black around, like it's the same color. So it would be eight and seven. [Terionna circles these on her array, circling the black-black flag first, and then the seven flags horizontally followed by the seven flags vertically (Figure 18).]

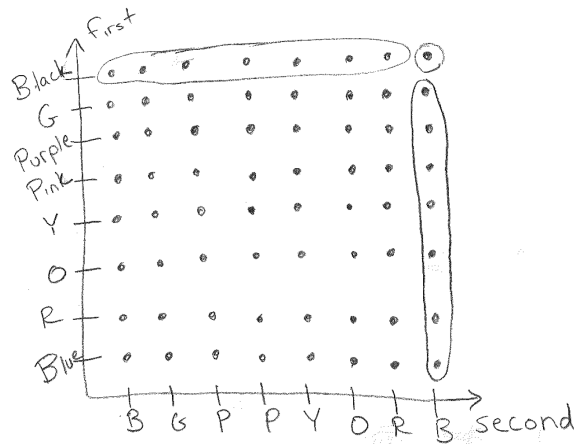


Figure 18. Terionna's array.⁹

Analysis of issue 3: Use of ordered pairs. Terionna established that there would be 15 flags with black in either the first or second position or both positions. She did so in the absence of looking at her written list, tree diagram, or array, which meant that she could not simply locate and count the ordered pairs in one of those representations. To start her solution, she first stated there would be eight flags with black in them. At the request of the interviewer, she generated a verbal list of these eight flags, saying the black-black flag last. After saying the black-black flag, she said, “wait, you are (“you are” meaning “she was”) just making black on the bottom and the other colors on top,” and adjusted her answer to 16 flags. I infer her adjustment from 8 to 16 flags came from verbally listing the black-black flag, which helped her realize that black could

⁹ This array is a replica of her original array. Her original had too many markings on it to reproduce it. Note that she labeled the vertical axis as representing the color in the first position in a pair and the horizontal axis as representing the color in the second position in the pair, which reverses the conventional way that ordered pairs are read.

also be a color in the top stripe of the flag. The interviewer asked her again to verbally list the 16 flags she was counting; she did, but again stopped after she said the black-black flag, which was only the eighth flag. Here she seemed to realize that for the black-black flag she could not switch the color in the first position with the color in the second position to produce a new flag, saying “you can’t switch black and black around.” Her realization enabled her to adjust the total number of flags again from 16 to 15.

To solve this problem, I infer Terionna used her concept of *ordered pairs*. I base this inference on her explanation of why she thought there would be sixteen flags. During this explanation, she said, “sixteen different groups of the black first and color, and then color and then black second.” This explanation indicated that she was not just envisioning switching the colors around in the flag; she was also relying on which position the black appeared in an ordered pair, a key criterion for the inference that a student was using ordered pairs. Moreover, I infer that Terionna engaged in a units coordination, which established a multiplicative relationship among a unit of one, a unit of eight units, and a unit of eight ordered pairs *in activity* (Figure 19). I make this inference because Terionna could reason about features of the eight ordered pairs without *always* having to create them individually. However, I considered her to produce this relationship *in activity* because she needed to produce *some* of the ordered pairs in order to establish the relationships. Establishing this relationship enabled her to reason that

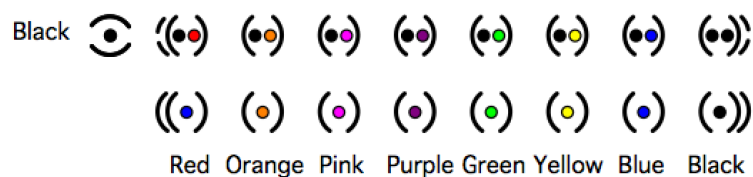


Figure 19. Multiplicative relationship among a unit of one, a unit of eight units, and a unit of eight pairs in activity

the black-black flag was contained in the flags that had black in the second position even in the absence of producing all of these flags, but to consider this issue she produced all flags with black in the first position. Figure 19 shows that she produced this multiplicative relationship in activity by using dotted parenthesis around the pairs (compare with Figure 16c).

Recall that only six of the MC2 students constructed ordered pairs. Of these students, I could attribute Terionna's way of operating to three of them. The other three MC2 students did *not* produce this way of operating. These students could locate ordered pairs in *pairs*, anticipating that for any pair they were given there was exactly one other pair that had the same colors in it. I considered this anticipation along with the operations I could infer about how they operated on a single composite unit (as shown with Shonda) to be sufficient to conclude that these MC2 students established a multiplicative relationship among a unit of one, a unit of one, and an ordered pair.

Discussion

Response to Research Questions: Identifying Domain Specific Contributions of the Study

Table 3 provides a summary of MC1 and MC2 students' solutions of APs from the Data Analysis section, and it compares this information to MC1 and MC2 students' solutions of CPPs. The individual cells of Table 3 provide responses to the first and second research questions, and a comparison of the rows in Table 3 provide a response to the third research question. I amplify these responses through a discussion of the three issues that I used to structure Data Analysis: (a) students' use of the odometer method in their solution of APs; (b) students' assimilation of APs using a single composite unit; and (c) students' construction of ordered pairs. Through this discussion I identify contributions of the study to the literature on combinatorial reasoning.

1 Table 3. Comparison of Students' Schemes for Solving APs and CPPs

	Differences in assimilatory structure of the scheme	Differences in the activity of the scheme	Difference in the results of the scheme
MC1 students	All three MC1 students assimilated APs using a single composite unit. In contrast, they assimilated CPPs using two composite units.	<ul style="list-style-type: none"> • For APs, all three MC1 students operated on the single composite unit by separating it into two parts (e.g., one and seven). They then used a pairing operation to produce pairs. • For CPPs, two of the MC1 students used a <i>disembedding</i> and pairing operation. By doing so, they established pairs as independent from, but related to the units of one they used to create them. • For APs, they did not use the odometer method to organize producing pairs. • For CPPs, they did use the odometer method. 	<ul style="list-style-type: none"> • For APs, two of the three students established ordered pairs. • For CPPs, they did not establish ordered pairs.
MC2 students	All eight MC2 students assimilated APs using a single composite unit. In contrast, they assimilated CPPs with two composite units.	<ul style="list-style-type: none"> • For APs, seven of the eight MC2 students <i>recursively</i> used their disembedding operation to solve APs: they disembedded a singleton unit from the composite unit they used in assimilation; and then they disembedded the singleton unit and first unit from the composite unit to create pairs. • For CPPs MC2 students used their disembedding operation, but did not use it <i>recursively</i> because they assimilated situations with two composite units. • For APs, six of the MC2 students used the odometer method to structure producing pairs. • For CPPs, all eight students used the odometer method. 	<ul style="list-style-type: none"> • For APs, six of the eight MC2 students constructed ordered pairs. • For CPPs, they did not establish ordered pairs. • For APs, of the six who constructed ordered pairs, three established a multiplicative relationship between a unit of one, a unit of one, and an ordered pair. These three students could anticipate finding ordered pairs in pairs in their lists, tree diagrams, or arrays. • For APs, the other three MC2 who constructed ordered pairs established a multiplicative relationship among a unit of one, a composite unit, and a unit of ordered pairs in activity. These three students were able to use ordered pairs to structure all flags with a particular color in the first or second position as a row and column of an array.

The findings from this study about MC1 and MC2 students' use of the odometer method, coupled with Author's (2014a) earlier findings, provides a trajectory for how students can re-construct and make advances in their use of this method. As students transition from solving CPPs to solving APs, students may move from: (a) using a consistent pattern for the items in the second position from column to column even though they do not use the *same order* for the items in the second position from column to column (what MC1 students did); to (b) using the same order for items in the second position from column to column, but not using the same order for the items in the first and second position (what the majority of MC2 students did); to (c) using the same order for the items in the second position from column to column *and* using the same order for items in the first and second position (what MC3 students did in Author, 2014a). This trajectory helps to identify one way the odometer method can become more advanced: Students can come to treat the order of the items in the first and second position as dependent on one another. Overall, this finding points to the fact that students may re-construct the odometer method across multiple kinds of combinatorics problems prior to constructing the method as a generalized method (Halani, 2012) that they can use to justify that they have produced all possible outcomes (Maher, et al., 2010).

The finding that MC1 and MC2 students in this study assimilated APs using a single composite unit, and as a result experienced a perturbation, confirmed Author's (2014a) earlier results with MC3 students. The prevalence of the perturbation among the students in the two studies helps to identify a key way that students experience APs as different from CPPs. One way to interpret this particular difference is that it is "in the problems;" APs refer to only a single set when they are stated (e.g., 8 colors) whereas CPPs refer to two sets when they are stated (e.g., 4 shirts and 3 pants). I do not interpret the difference in this way because it is possible to

interpret an AP like the Flag Problem as involving two composite units; a student could anticipate, for example, that there are eight possible colors for the top stripe and eight possible colors for the bottom stripe prior to producing any flags. In fact, one MC3 student in Author (2014a) learned to interpret APs in this way. In the Implications Section, I provide an analysis of why interpreting APs in this way can be mathematically significant. This analysis provides additional rationale for why it is important to understand how students resolve this commonly experienced perturbation.

One interesting difference between this study and Author's (2014a) earlier study was that in this study there was a more explicit design to promote the construction of ordered pairs as students transitioned from CPPs to APs. This design allowed for the identification of three criteria for defining when a student has constructed ordered pairs. Establishing such criteria is significant given that researchers have found: (a) that students often do not consider counting ordered pairs in combinatorial contexts (e.g., Author, 2014a; Fischbein & Gazit, 1988; Piaget & Inhelder, 1975;), and (b) that students' lack of construction of ordered pairs often adversely impacts how they reason about probabilistic situations (Fischbein, Nello, & Marino, 1991; Morsanyi, Handley, & Serpell, 2013; Rubel, 2007). The findings from these earlier studies suggest that even when students do consider two outcomes to be different (e.g., the red-blue and blue-red flag) this alone is not sufficient for students to reason with ordered pairs in normative ways, which is why in this study I considered this a necessary, but not sufficient criterion to infer the construction of ordered pairs.

Response to Research Questions: Identifying Theoretical Contributions of the Study

Current models of student reasoning have not yet carefully traced out whether and when students establish relationships among one and two-dimensional units (Author, 2016; Battista,

2007; Smith & Barrett, 2018). The findings from this study provide one example of models of student reasoning that explicitly attend to this issue in combinatorial contexts. Doing so uses a mathematical distinction that Vergnaud (1983) made to provide more nuanced psychological models of students' multiplicative reasoning, extending earlier work that has attended to unit structures (e.g., Steffe, 1992, 1994; Hackenberg, 2007, 2010), but which has not explicitly modeled whether and when students establish relationships among one and two-dimensional units. A discussion of the differences in operations that students used to solve CPPs and APs provides insight into how students established and conceptualized these relationships differently in the two problem contexts.

To solve CPPs, MC1 students used a disembedding operation in concert with their pairing operation. Using this combination of operations enables a student to establish one two-dimensional unit as independent of, but related to, the two one-dimensional units that comprise it (Author, in press). MC1 students' operated differently in their solution of APs; they used a pairing operation, but *not* a disembedding operation. Operating in this way meant that in their solution of APs they did establish a two-dimensional unit (i.e., a pair), but they did not retain the two one-dimensional units that comprised the pair. This change in how they operated was in part a result of assimilating APs using a single composite unit, which meant to retain the two one-dimensional units as part of the situation would involve a recursive use of their disembedding operation. There was not sufficient evidence to conclude that MC1 students operated in this way. Providing models of student reasoning that account for when they do and do not establish this relationship is important given researchers' findings that students do not easily establish two-dimensional units as the product of two one-dimensional units (Nunes, Light, & Mason, 1993; Outhred & Mitchelmore, 2000).

MC2 students established similar multiplicative relationships in their solution of APs and CPPs; three MC2 students had interiorized a multiplicative relationship among a unit of one, a unit of one, and an ordered pair and three MC2 students produced a multiplicative relationship among a unit of one, a unit of eight units, and a unit of eight ordered pairs in activity. The primary difference was that in APs to establish these multiplicative relationships they used their disembedding operation recursively. One way to think about the importance of using one's disembedding operation recursively is that it means that a student creates the one-dimensional units from the same material rather than conceptualizing the one-dimensional units as qualitatively distinct (e.g., as shirts and pants). When students operate in this way in their solution of APs, they can make a qualitative distinction between the one-dimensional units (e.g. colors for the first stripe and colors for the second stripe), but the qualitative distinction results from them imposing it on what they conceive of as one-dimensional units that do not initially have a qualitative distinction. Providing models that allow researchers to make this kind of inference is important given that students frequently treat qualitative labels (e.g., "length" and "width") as if they have an impact on the quantitative units of measure (Battista, 2007).

Limitations of the Study

One of the central limitations of the current study is that it was conducted over a short period of time. Therefore, it gives a small snapshot about what students operating with different multiplicative concepts could do, but does not provide information about what learning is possible over an extended period of time. Maher & colleagues (2010) have provided one of the few longitudinal studies in the domain of combinatorics. Other longitudinal studies are needed. A second limitation of the study is that the results are tentative in the sense that researchers have not made explicit models of the relationships students establish among one and two-dimensional

units in combinatorial contexts. The process of making these models is in its early stages, and continued work is needed to clarify and refine the models, while also considering how they could be integrated into what we already know about students' multiplicative (e.g., Greer, 1992; Verschaffel, Greer, & De Corte, 2007), geometric (e.g., Battista, 2007), and measurement based reasoning (e.g., Smith & Barrett, 2018). Both limitations point to directions for future research.

Implications and Directions for Future Research

One primary reason for attending to students' transition between CPPs and APs has to do with a mathematical difference between these problems. The mathematical difference stems from Vergnaud's (1983) idea that product of measures problems involve what he called a double proportion. In the case of CPPs, Vergnaud's double proportion means that, for example, the total number of outfits in the Outfits Problem is proportional to both the number of shirts and the number of pants. Therefore, if the number of pants is tripled, then the total number of outfits is tripled (Figure 20a), and similarly if the number of shirts is tripled, then the total number of outfits is tripled (Figure 20b). In the case of CPPs, these two proportional relationships can be considered independent of each other—changing the number of pants can be done independently from changing the number of shirts.

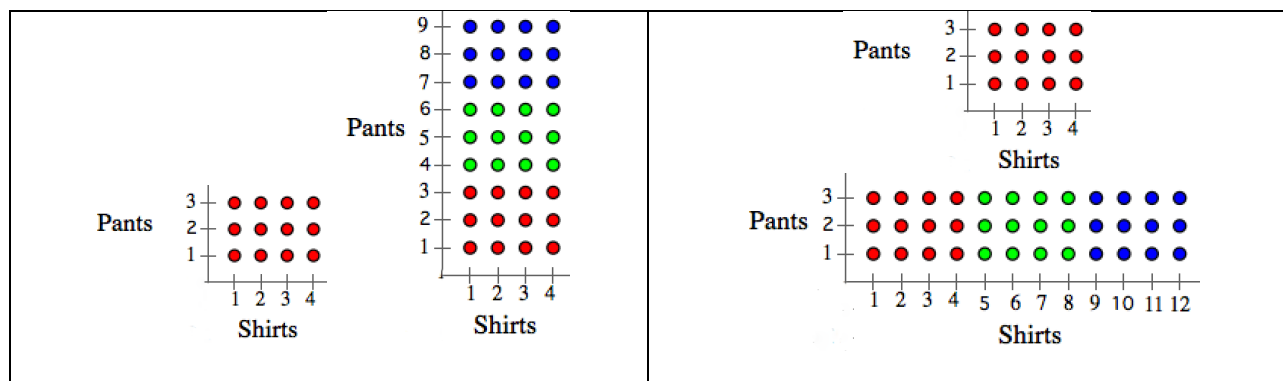


Figure 20a (left) & 20b (right): Arrays to represent the multiplicative relationships

On the other hand, APs like the Flag Problem can be seen as a special case of the double proportion property (Author, 2013); the total number of two striped flags is proportional to the number of colors for the first stripe and the number of colors for the second stripe, however, the number of possible colors for the second stripe can be interpreted as *dependent* on the number of possible colors for the first stripe. So, for example, doubling the number of colors in the Flag Problem means that the number of colors for the first stripe is doubled *and* the number of colors for the second stripe is doubled, which produces *four* times the total number of two striped flags. This lack of independence of the proportions means that the double proportion functions as a direct *square* proportion: a change in the number of colors is related to a change in the number of two-striped flags through a quadratic relationship (Figure 21).

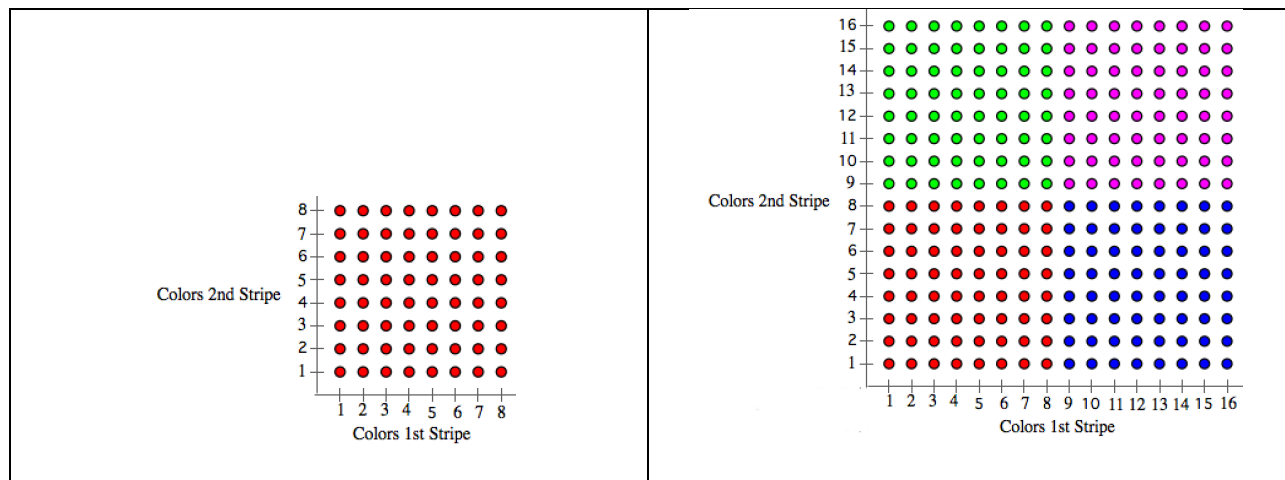


Figure 21. Array for the Flag Problem

This mathematical distinction suggests that the way students understand the relationship in APs between, for example, the number of colors for the first stripe and the number of colors for the second stripe has important mathematical implications; this relationship is the basis for how students might establish quadratic relationships in their solution of APs. The relationship students establish is a function of what operations they use to resolve the perturbation they experience when they assimilate APs using a single composite unit. Thus, one implication of this

study, and direction for future research, is to understand how solving APs might support students to establish quadratic relationships. Understanding how students establish quadratic relationships is critical because of how pervasive it is for students to reason as if *all* situations involve linear relationships (e.g., Greer, 2010; Van Dooren, De Bock, Janssens, & Verschaffel 2008).

Conclusion

The increased focus in standards and curricula on combinatorics, specifically, and discrete mathematics, more broadly, accompanied with research indicating that this domain is quite challenging for students (e.g., Fischbein & Gazit, 1988; Halani, 2012), suggests a need for researchers to continue to work on developing models of student reasoning that can be used to support instructional decision making (e.g., English, 1991, 1993; Lockwood & Swinyard, 2016). The findings from this article provide one example of this kind of model. As these models are developed, refined, and extended, they can serve as powerful instruments to support teachers and researchers to engage students in productive interactions (Author, 2014b). Moreover, attending to how students who use different multiplicative concepts reason about similar situations can serve as a basis for researchers and teachers to differentiate instruction in ways that are rooted in differences in student reasoning (Hackenberg, Creager, & Eker, & Lee, 2016). Differentiation of instruction helps researchers and teachers make informed decisions about how to adjust to students who bring different resources to the classroom.

References

Author (2009).

Author (2013).

Author (2014a).

Author (2014b).

Author (2016).

Author (in press).

Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32(2), 181-199.

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 843-908). Charlotte, NC: Information Age Publishing.

Behr, M., Harel, G., Post, T., & Lesh, R. (1994). Units of quantity. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 121-176). Albany, New York: State University of New York Press.

Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In A. E. Kelly & L. R. (Eds.), *Handbook of research design in mathematics and science education* (pp. 547-589). Mahwah, NJ: Erlbaum.

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture-driven research design. In A. E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 231-265). Mahwah, NJ: Lawrence Erlbaum Associates.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (third ed.). Thousand Oaks, CA: Sage

Publications.

- Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning*, 6(1), 15-36.
- English, L. D. (1991). Young children's combinatoric strategies. *Educational Studies in Mathematics*, 22, 451-474.
- English, L. D. (1993). Children's strategies for solving two- and three-dimensional combinatorial problems. *Journal for Research in Mathematics Education*, 24, 255-273.
- English, L. D. (1996). Children's construction of mathematical knowledge in solving novel isomorphic problems in concrete and written form. *Journal of Mathematical Behavior*, 15, 81-112.
- English, L. D. (1999). Assessing for structural understanding in children's combinatorial problem solving. *Focus on Learning Problems in Mathematics*, 21(4), 63-83.
- Fischbein, E., & Gazit, A. (1988). The combinatorial solving capacity in children and adolescents. *Zentralblatt fur Didaktik der Mathematik*, 5, 193-197.
- Fischbein, E., Nello, M. S., & Marino, M.S. (1991). Factors affecting probabilistic judgments in children and adolescents. *Educational Studies in Mathematics*, 22(6), 523-549.
- Greer, B. (1992) Multiplication and division as models of situations. In Grouws, D.A. (Ed.) *Handbook of Research on Mathematics Teaching and Learning*, pp. 276-295. New York: Macmillan.
- Greer, B. (2010). Overview of papers: Why is linear thinking so dominant? *Mathematical Thinking and Learning*, 12(1), 109-115.
- Hackenberg, A. J. (2007). Units coordination and the construction of improper fractions: A revision of the splitting hypothesis. *Journal of Mathematical Behavior*, 26, 27-47.

- Hackenberg, A. J. (2010). Students' reasoning with reversible multiplicative relationships. *Cognition and Instruction*, 28(4), 1-50.
- Hackenberg, A. J., Creager, M., Eker, A., & Lee, M. L. (2016, April). *Understanding how to differentiate instruction for middle school students*. Paper presentation at the Research Conference of the NCTM, San Francisco, CA.
- Halani, A. (2012). Students' ways of thinking about enumerative combinatorics solution sets: The odometer category. In the *Electronic Proceedings for the Fifteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education*. (pp. 59-68) Portland, OR: Portland State University.
- Kapur, J. N. (1970). Combinatorial analysis and school mathematics. *Educational Studies in Mathematics*, 3(1), 111-127.
- Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 909-955). Charlotte, NC: Information Age Publishing.
- Lappan, G., Fey, J.T., Fitzgerald, W.M., Friel, S.N., & Phillips, E.D. (2002). *Clever counting*. Upper Saddle River, NJ: Prentice Hill.
- Lobato, J. (2008). Research methods for alternative approaches to transfer: Implications for design experiments. *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*, 167-194.
- Lockwood, E. (2011). Student connections among counting problems: An exploration using actor-oriented transfer. *Educational Studies in Mathematics*, 78(3), 307-322.
- Lockwood, E. (2013). A model of students' combinatorial thinking. *The Journal of Mathematical Behavior*, 32(2), 251-265.

- Lockwood, E., & Gibson, B. (2016). Combinatorial tasks and outcome listing: Examining productive listing among undergraduate students. *Education Studies in Mathematics*, 91(2), 247-270.
- Lockwood, E. & Swinyard, C.A. (2016). An introductory set of activities designed to facilitate successful combinatorial enumeration for undergraduate students. *Problems, resources, and issues in mathematics undergraduate studies*, 26(10), 889-904.
- Maher, C.A., & Martino, A. (1996). The development of the idea of mathematical proof. *Journal for Research in Mathematics Education*, 27(2), 194-214.
- Maher, C. A., Powell, A. B., & Uptegrove, E. B. (Eds.). (2010). *Combinatorics and reasoning: Representing, justifying and building isomorphisms* (Vol. 47). Springer Science & Business Media.
- Maher, C.A., & Yankewitz, D. (2010). Representations as tools for building arguments. In C.A. Maher, A.B. Powell, & E.B. Uptegrove (Eds.), *Combinatorics and reasoning: Representing, justifying, and building isomorphisms* (pp. 17-25). New York: Springer.
- Martin, G. E. (2001). *The art of enumerative combinatorics*. New York: Springer.
- Mathison, S. (1988). Why triangulate? *Educational Researcher*, 17(2), 13-17.
- Morsanyi, K., Handley, S. J., & Serpell, S. (2013). Making heads or tails of probability: An experiment with random generators. *British Journal of Educational Psychological Society*, 83, 379-395.
- Mulligan, J., & Mitchelmore, M. (1997). Young children's intuitive models of multiplication and division. *Journal for Research in Mathematics Education*, 28(3), 309-330.
- National Council of Teachers of Mathematics (2000). *Principles and standards of school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Governors Association for Best Practices, Council of Chief State School Officers (2010).

Common core state standards mathematics. Washington D.C.: National Governors Association for Best Practices, Council of Chief State School Officers.

Nunes, T., & Bryant, P. (1996). The progress to multiplication and division. *Children doing mathematics* (pp. 142-201). Oxford, UK: Blackwell Publishers.

Nunes, T., Light, P., Mason, J. (1993) Tools for thought: The measurement of length and area. *Learning and Instruction*, 3, 39-54.

Outhred, L. (1996). Children's drawings of multiplicative structures: Cartesian products and area. In J. Mulligan & M. Mitchelmore (Eds.), *Children's number learning* (pp. 185-204). Adelaide, Australia: Australian Association of Mathematics Teachers for the Mathematics Education Research Group of Australasia.

Outhred, L. & Mitchelmore, M. (2000). Young children's intuitive understanding of rectangular area measurement. *Journal of Research in Mathematics Education*, 31(2), 144-167.

Piaget, J. (1970). *Genetic Epistemology* (E. Duckworth, Trans.). New York: W.W. Norton & Company Inc.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children* (L. J. Leake, P. Burrell & H. D. Fishbein, Trans.). New York: W.W. Norton & Company.

Rubel, L. H., (2007). Middle school and high school students' probabilistic reasoning on coin tasks. *Journal for Research in Mathematics Education*, 38(5), 531-556.

Saldana, J. (2013). *The Coding Manual for Qualitative Researchers*. London: Sage Publications Ltd.

Senk, S. L., Viktora, S. S., Usiskin, Z., Ahbel, N. P., Highstone, V., Witonsky, D., et al. (1998). *Functions, statistics, and trigonometry* (2nd Edition). Glenview, IL: Scott Foresman Addison Wesley.

- Simon, M. A., & Blume, G.W. (1994). Building and understanding multiplicative relationships: A study of prospective elementary teachers. *Journal for Research in Mathematics Education*, 25(5), 472-494.
- Smith, J.P., III, & Barrett, J.E. (2018). Learning and teaching measurement: Coordinating quantity and number. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 355-385). Reston VA: National Council of Teachers of Mathematics.
- Steffe, L. P. (1992). Schemes of action and operation involving composite units. *Learning and individual difference*, 4(3), 259-309.
- Steffe, L. P. (1994). Children's multiplying schemes. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 1-41). Albany, New York: State University of New York Press.
- Steffe, L. P. (2007, April). *Problems in mathematics education*. Paper presented for the Senior Scholar Award of the Special Interest Group for Research in Mathematics Education (SIG-RME) at the annual conference of the American Educational Research Association in Chicago, Illinois.
- Steffe, L.P., & Olive, J. (2010). *Children's fractional knowledge*. Boston, MA: Springer.
- Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 267-307). Hillsdale, NJ: Erlbaum.
- Steffe, L. P., von Glasersfeld, E., Richards, J., & Cobb, P. (1983). *Children's counting types: Philosophy, theory, and application*. New York City: Praeger Publishers.
- Tucker, A. (2002). *Applied combinatorics* (4th ed.). New York: John Wiley & Sons.
- Ulrich, C. (2015). Stages in Constructing and Coordinating Units Additively and Multiplicatively (Part 1). *For the Learning of Mathematics*, 35(3), 2-7.

- Ulrich, C. (2016). Stages in Constructing and Coordinating Units Additively and Multiplicatively (Part 2). *For the Learning of Mathematics*, 36(1), 34-39.
- Ulrich, K., Author, E.S., Hackenberg, A.J., Norton, A.N. (2014). Units coordination: An example of the utility of radical constructivist thought in education. *Constructivist Foundations*, 9(3), 328-339.
- Van Dooren, W., De Bock, D., Janssens, D., & Verschaffel, L. (2008). The linear imperative: An inventory and conceptual analysis of students' overuse of linearity. *Journal for Research in Mathematics Education*, 39(3), 311-342.
- Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127-174). New York: Academic Press.
- Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 557-628). Charlotte, NC: Information Age Publishing.
- Von Glasersfeld, E. (1982). Subitizing: The role of figural patterns in the development of numerical concepts. *Archives de Psychologie*, 50, 191-218.
- Von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning* (Vol. 6). Washington D.C.: The Falmer Press.

Appendix A: Selection Interview Tasks

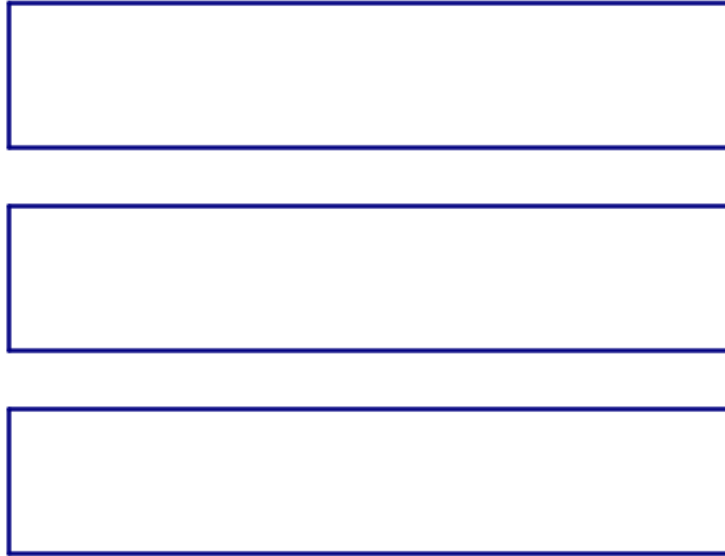
- 1)
 - a. The 7th grade classrooms has 4 rows of desks with 14 desks in each row. How many desks are in a 7th grade classroom?
 - b. Suppose that 28 desks and 1 row are added to the classroom. How many rows and desk would there be?

- 2) You and 11 friends go to the movies (12 people total). You and your friends fill the first 3 rows of the theater. Determine the number of seats in the theater if there are 13 rows. (Again all rows have the same number of seats.)

- 3) The 8th grade is taking 6 buses on a field trip. There are 96 students in the 8th grade. Use the picture below to show how many students will be on each bus (assume there are an equal number of students on each bus).

- 4)
 - a. A candy factory puts 10 candies in each package and puts 10 packages in each box. The factory currently has 3 boxes, 4 packages, and 7 candies. How many total candies do they have?
 - b. You buy 534 candies from the candy factory. How many packages of candy do you get?

- 5) Five people want to share the three identical candy bars. Use the bars below to show how you would accomplish this goal.



- 6) Jose buys some rope that is 65 centimeters long. Jose's rope is 5 times the size of the piece of rope you bought. Below is a picture of Jose's rope. Use the picture to help you determine the length of your rope.



- 7) Stephanie checks out 14 books on spiders from the library to prepare for her science fair project. The number of books Stephanie checked out is $\frac{2}{7}$ of the total number of books the library owns on spiders. How many total books does the library own on spiders?

Appendix B: Interview 1 Tasks

- 1) You have 4 shirts and 3 pairs of pants. An outfit is 1 shirt and 1 pair of pants. Draw a picture to help you determine how many total outfits you can make.
- 2) Subway has 6 kinds of bread and 7 different kinds of meat. You can make a sandwich by choosing one kind of meat and one kind of bread. Draw a picture to help you determine how many total sandwiches you can order.
- 3) You have the Ace thru 7 of spades and your friend has the Ace thru 7 of hearts. A two-card hand consists of one card from your friend and one card from you (e.g., the two of spades and the two of hearts would be 1 two-card hand). How many different two-card hands could you make?
- 4) You and a partner each flip a coin and record the result of each of your coin flips (e.g., tails-tails would mean you each got tails on a flip). How many different outcomes could you record?
- 5) You are designing a flag for a recently discovered country. The flag has two stripes, each of which can be filled with a color. You have 8 different colors to choose from. Illustrate the total number of flags you could make.
- 6) A volleyball league has 10 teams. Each team wants to play each other once at home and once away. How many total games would there be?
- 7) You have the ace thru king of spades. To create a two card hand, you draw a card, replace it, and then draw a second card. How many possible two card hands could you make?
- 8) You are creating a two-character password from the letters A through N. How many different passwords could you create?